NANO MICRO small

Supporting Information

for *Small*, DOI: 10.1002/smll.201501178

Critical Points in Tumorigenesis: A Carcinogen-Initiated Phase Transition Analyzed via Single-Cell Proteomics

*Suresh Kumar Poovathingal, Nataly Kravchenko-Balasha, Young Shik Shin, Raphael David Levine, and James R. Heath**

Supporting Information

Critical Points in Tumorigenesis: A Carcinogen Initiated Phase Transition Analyzed via Single Cell Proteomics

*Suresh Kumar Poovathingal, Nataly Kravchenko-Balasha, Young Shik Shin, R. D. Levine, and James R. Heath**

1. Supplementary Text:

ST1. SCBC fabrication and operation:

Antibody microarray which is used to quantify the protein concentration at single cell level is based on a DNA encoded antibody technology.[1, 2] Antibody microarray comprises of spatially distinct array of single stranded DNA, which is flow-patterned on an poly-l-lysine glass substrate using high-density PDMS template (details in Supporting Information ST3 and ST4). Antibody array is prepared by hybridizing the antibody-DNA conjugate, just prior to cell loading and lysis process. The microarray has the capability of multiplexed measurement up to twenty different proteins. Details of the microarray fabrication is detailed in Supporting Information ST3.

SCBC device which encapsulates and quantifies the protein concentration at single cell level is based on two-layer soft lithography (details in Supporting Information ST4). The two layer PDMS device was fabricated as described elsewhere [1, 2]. The protocol for fabrication and operation of the SCBC device is based on previously published work,[2] however with several modification to suite the cell handling and loading of the present system in consideration (detailed protocol in Supporting Information ST5). The SCBC assay measurements were calibrated to yield absolute protein copy numbers (Supporting Information ST6 and ST7).

ST2. DEAL conjugate synthesis:

Detailed protocol for the DNA-Antibody conjugation can found in Protein-Oligo conjugation kit (Solulink). Capture antibodies are desalted and buffer exchanged to pH 7.4 using Zeba protein desalting column (Pierce). Succinimidyl 4-hydrazinonicotinate acetone hydrazine (SANH, Solulink) is mixed in N,N dimethylformamide (DMF) was added to the desalted antibody solution in molar excess; (300:1 of SANH to antibody). Separately, succinimidyl 4- formylbenzoate (SFB) is mixed with DMF and was added to 5'-aminated oligomer in PBS in a molar excess of 16-fold (16:1; of SFB to DNA). These are allowed to react separately for 4 hrs at room temperature. Excess unreacted SANH and SFB are removed from both the samples and they are buffer-exchanged with Zeba desalting

column to pH 6.0 using citrate buffer. The DNA and antibody was combined and allowed to react overnight to form the conjugate. Non-reacted DNA and antibody is excluded using Pharmacia 200 gel filtration column (GE). The conjugates were then concentrated to ~0.5 mg/ml using Amicon Ultra-4 centrifugal filter units (Millipore 10kDa) and kept for long-term storage at 4^0C.

ST3. DNA barcode fabrication:

Fabrication of the Barcode slides have been described in extensive details elsewhere.[1, 2] Here, only modifications of that protocol are highlighted. A polydimethyl siloxane (PDMS) elastomer substrate is prepared by mixing the base elastomer of SYLGARD® 184 Silicone with the binding agent in a mixing ratio of 10:1. A flow patterning silicon mold is used to produce a PDMS microfluidic layer for flow patterning the barcodes onto polylysine glass slides (VWR). The patterning PDMS substrate comprises of 20 parallel channels. For the present work only 12 channels were utilized. The initial patterns are of unique ssDNA oligomers. After flow patterning is complete, the flow patterning PDMS layer is removed from the glass substrate, which can then be mated to separate PDMS layers to form the SCBC. Of the 12 channels, 11 of them were used for the protein quantification and one was used as the reference marker (see Supplementary Table S1 for a list of all ssDNA and antibody reagents used).

ST4. SCBC Fabrication:

The PDMS integrated microfluidic chip for the single cell experiments were fabricated using two-layer soft lithography. Details of SCBC fabrication and operation have been previously published.[2] A change in the published protocol was that both the control and the flow layer were fabricated using SYLGARD® 184 Silicone elastomer. The control layer and the flow layer were initially molded separately using an elastomer mix of 7:1 (A:B) and 20:1 (A:B), respectively. Thin flow layer is made by spin coating the PDMS mix over the silicon wafer at a speed of 2000 rpm for 60 seconds. Both the layers are cured separately at 80^0C for 15 minutes. After curing, individual PDMS substrates are cooled to room temperature, and the control substrate was cut out and aligned with the flow layer. The aligned PDMS substrates were further cured at 80^oC for 75 minutes. This curing ensured strong boding between the control and flow layers. The bonded PDMS substrate was then peeled off the surface of the silicon mold, and holes were punched for input and output access. Finally, the DNA barcoded poly-lysine slide is properly aligned and thermally bonded to the PDMS chip at 80^0C for 2 hours to produce an assembled device. The device is designed so that each of ~300 microchambers within a given SCBC contains a full 12-element barcode array.

ST5. SCBC Operational details:

The protocols for using the SCBC device for profiling the intracellular proteins from single cells are described below:

- The microfluidic channels are blocked with 1.5% BSA solution for one hour to reduce non-specific binding.
- Just prior to use, the ssDNA barcoded slide is converted to a capture antibody microarray through the DEAL technique, in which a cocktail of 11 antibody-ssDNA' conjugates are flowed over the DNA patterned poly-lysine slide for 60 minutes at 37^0C. Simultaneously, concentrated cell-lysis buffer with protease and phosphatase inhibitors were loaded to lysis buffer channel (channel 2 in Figure S1b). The flow channel (channel 1 in Figure S1b) and the lysis buffer channel (channel 2 in Figure S1b) are isolated by control valve 4 (5 in Figure S1b), which is operated under 18 psi of pneumatic pressure. In order to remove the unconjugated DNA-antibody conjugates and to introduce physiological condition before loading the MCF-10F cells, the micro channels were flushed with the complete growth medium for 10 minutes.
- MCF-10F cell samples are serum starved for 24 hours prior to loading into the SCBC device. MCF-10F cells are trypsinized using 0.05% Trypsin-EDTA solution. When the cells are about to detach (attain circular shape), the activity of Trypsin-EDTA is diluted with the solution of excess serum starved medium. The cells are pelleted by centrifuging the cell suspension at 200g for 5 minutes. Prior to loading, cells were activated by suspending the pellet in serum starved medium dosed with 100 ng/ml of EGF and 100 ng/ml of Insulin and cell density in the medium is maintained ~1000 cells/ml. Within 30 minutes of growth factor treatment, the cells are loaded into the cell chambers and subsequently lysed. After loading the cells into the microchannels, valve 3 (6 in Figure S1b) is closed to compartmentalize the channels into 310 micro chambers. Images of each chamber were recorded using a CCD camera and used for tabulating cell count per micro chamber.
- The SCBC chip is placed on ice and a the separating valve (5 in Figure S1b) is opened to allow the lysis buffer to diffuse from lysis channel (2 in Figure S1b) to the cell chambers (1 in Figure S1b). The valve is opened for ~20 minutes to allow complete diffusion of lysis buffer into the microchamber. After which, the valve is closed and the SCBC chip is incubated on ice for ~40 mins for complete cell lysis and diffusion. The chip was further incubated at room temperature for 2 hours to complete the capture of specific intracellular proteins onto the antibody barcode microarray within the respective microchambers. After the incubation, unbound cell-lysate and debris is flushed out using flowing 1.5% BSA solution for 10 minutes.

- A mixture of biotinylated-antibodies is prepared by mixing the antibodies at recommended ELISA concentration in 1.5% BSA solution. The detection antibody was flown through the channel for 60 minutes at room temperature to complete the immune-sandwich assay. Unbound detection antibody was flushed out of the SCBC device by washing the entire device with 1.5% BSA solution for 10 minutes.
- Streptavidin-Cy5 (eBioscience, conc: 2µg/ml) and marker reference, Cy3-labelled complementary ssDNA (M-M' at conc: 25nM) was prepared in 1.5% BSA solution and was flown through the channels for 60 minutes at room temperature. Unbound labels are flushed out of the system by washing the microchannels with 1.5% BSA for ~45 minutes.
- After washing, the PDMS substrate is peeled of the glass surface and surface of the glass slide is washed by sequentially dipping the slide in 0.5X PBS, deionized water and finally dried under the flowing N_2 stream.
- Optical readout: The slide was scanned by an Axon GenePix 4400A (Molecular Devices) at the laser power of 80% (635 nm) and 10% (532 nm) and at a resolution of 2.5 µm. Average fluorescence readout from all the cell chambers were further extracted using a custom built MATLAB code. Custom software then translates the fluorescence readout into a table that contains the microchamber address, the number of cells in that microchamber, and the copy numbers for each molecule assay

ST6. Calibration curve:

A useful requirement for applying Surprisal analysis in characterizing the critical transition process is the absolute copy number determination from the fluorescence readout. These absolute quantification or calibration was performed using SCBC device and under the exact same condition as was done with the single-cell proteomic assay described above, except that standard proteins were used instead of cells in the single cell chambers. A mixture of standard recombinant proteins (detail of the recombinant proteins in Supplementary table S1) for all the proteins in the assay panel was serially diluted in 1× PBS and flowed into the SCBC microchannels. Fluorescence readouts from different protein concentration was used to generate the calibration curves. Since the volume of microchambers are known, the calibration plot enables to determine the relation between fluorescence intensity and the absolute copy number of the protein being assayed, under the assumption that the recombinant proteins have same physical and chemical attributes, compared to their counterparts in the actual cell.

ST7. Absolute protein copy-number quantification:

The SCBC data is based on fluorescence units, and in order to covert the fluorescence measure to absolute protein copy number, the calibration plots were used (Figure 1). To fit the calibration data, five parameter regression function was used.

$$y = A_2 + \frac{A_1 - A_2}{\left[1 + \left(\frac{x}{x_0}\right)^p\right]^q} \quad (4)$$

Where, A_1, is the estimated response at zero concentration; A_2, is the estimated response at infinite concentration; *p*, slope factor of the response curve; *q*, asymmetry factor and x_0, mid-range concentration value. Statistical output from the single cell analysis is the input for the theory as a series of tables, one table per time point for each carcinogen used. Each column represents a measured protein level, while each row represents an individual cell.

Error Consideration: Error analysis related to Single cell proteomic measurement in SCBC has been extensively considered and discussed in previous works[1]. Briefly, it was previously shown that the location of a cell within a microchamber does not significantly contribute towards the measurement error. Furthermore, it was also demonstrated that the variation of the protein level is dominated by biological noise (cell-to-cell heterogeneity) rather than the experimental noise [2] and that the experimental error in measuring the level of a given protein is typically <10%.

ST8. Cell Lines and Reagents: MCF10F cell line (ATCC) was maintained and passaged in complete growth medium (1:1 mixture of DMEM and HAM's F12 and 5% horse serum. The medium is supplemented with mitogenic factor such as: 20ng/ml of epidermal growth factor, 500 ng/ml Hydrocortisone, 100ng/ml of Cholera Toxin, 10µg/ml of insulin.) as described elsewhere.[3] Cell culture medium was also supplemented with 100 U/ml penicillin and 100µg/ml of streptomycin and was maintained in 5% CO_2 at 37^0C. Aqueous stock solution of Benzo(a)pyrene (B[a]P) was prepared in dimthysulfoxide (DMSO). 1µM carcinogen dosage medium was prepared by dosing the respective stock solutions of the carcinogen in the growth medium.

ST9. Inducing carcinogenesis: As shown in Fig. 1 of the article, carcinogen treatment comprised of three treatment cycles: 24 hrs after splitting the cells into fresh growth medium, the growth medium along with the unattached cells are removed and replaced with the medium containing carcinogen. 48 hours post exposure, carcinogen dosed medium was replaced with fresh medium. The gap between successive exposures was 14 days. All the samples were passaged regularly at 3-4 days interval. All along the treatment and post-treatment, the cells were maintained in 5% CO_2 at 37^0C. After the exposure of the cells to carcinogens, phenotypic assays were performed to detect the

acquiring of cancer related properties in cells that survived the carcinogen treatment: Colonogenic assay for estimating the proliferative capacity under limited dilution, reduced dependence on growth factors and anchorage independent growth.

ST10. Colonogenic assays:

*Colonogenic assay in normal growth medium***:** This assay is a conventional method used to study the cyto-toxicity and proliferative capacity of mammalian cells growing in monolayer, attached to a surface[4] After the cell-split ~5000 cells were seeded into plates with complete growth medium. Control cells which haven't yet undergone carcinogenic transformation have relatively poor colony forming efficiency. Cells were maintained at 5% CO_2 at 37^0C incubation. After 8-days of incubation, colonies having size >100µm was quantified. Colony counting and image analysis was performed using ImageJ (NIH, Bathesda).

Colonogenic assay for reduced dependence on growth factor **:** After splitting the cells, ~20,000 cells were seeded into plates containing low-mitogen (LM) medium (1:1 mixture of DMEM/F12, 2ng/ml cholera endotoxin, 200ng/ml insulin, 10ng/ml hydrocortisol, 0.4 ng/ml EGF and 0.1% horse serum) containing reduced serum and mitotic additives to 2% (v/v) of the concentration formulated in the complete growth medium.[5] Cells were maintained at 5% CO_2 at 37^0C incubation. Growing cell colonies that reached a diameter of > 20µm in LM medium after 10 days of seeding, were identified as clones acquiring reduced dependence on growth factor. Colony counting and image analysis was performed using ImageJ (NIH, Bathesda).

*Colonogenic Assay for Anchorage independence***:** Base layer consisted of 2% low-gelling Select agarose (Invitrogen) in complete growth medium. Middle layer comprised of 0.5% agarose in a mixture of complete growth medium, warmed to ~37^0C, was seeded with ~20,000 cells per well and this was poured on to well-set bottom agarose layer.[5] The middle gel layer was also allowed to solidify at 37^0C. After which, fresh growth medium was topped up on the middle layer. Soft-agar cultures were maintained at 5% CO_2 at 37^0C. Colony formation and growth was observed microscopically and the cultures were incubated at 37^0C for 21 days. Fresh media was supplemented at a regular interval of ~3 days. After 21 days in culture, colonies having size >50µm was quantified. Colony counting and image analysis was performed using ImageJ (NIH, Bathesda). In all the colonogenic assays, the colonies were stained with crystal violet prior to counting.

ST11. Surprisal analysis:

An early application to average transcription levels and functional proteins is presented in [6-10]. It is extended in this article to quantify the changes of protein copy numbers in individual cells. In such an

approach each sampled cell is regarded as a typical representative of a large number of cells from the bulk culture.

We assume that cellular non-equilibrium system is in a state of maximal entropy subject to constraints at every time point during the processes of carcinogenesis. The state of the maximal entropy in biological cells is the steady state, a state without any constraints on the biological networks. Biological steady state is a state in which an internal biological process is maintained at a relatively constant level and intermediates are formed and consumed at equal rates.[11] For experiments at constant temperature (such as 37^0, typical in biology), an entropy maximum is equivalent to a free energy minimum, and the state of lowest free energy is the steady state. The constraints are detected and quantified by identifying how the protein expression levels respond to those constraints. For each measured proteins the extent of participation of the protein in the biological process responding to constraints is defined.

The experimentally determined values, $X_i(cell,t)$, of the functional proteins in every cell as a function of time are used to identify the constraints. Constraints prevent the entropy from reaching the global possible maximum, and thereby generate deviations from the steady state. We seek to define the protein concentrations in every cell , $X_i^O(cell,t)$ at the steady state-the state in order to build protein distributions at the biological global maximum of the entropy-and compare it to the experimental protein level distributions.

At every time point in course of carcinogen treatment the most stable state is the state of minimal free energy while deviations thereof necessarily have a higher free energy. We characterize the expression levels of proteins at a given point in time we use surprisal analysis as summarized in Equation (1) of the main text.

To obtain Equation (1) in the main text, we relate protein concentrations to the chemical potential (under constant temperature and pressure) using the fundamental physical chemical relations:[12]

$$\underbrace{\mu_i}_{\text{free energy of protein } i} = \underbrace{\mu_i^{ss}}_{\text{standard free energy}} + kT \ln X_i(cell,t)$$

$$\underbrace{\mu_i^{O}}_{\substack{\text{free energy of protein } i \\ \text{at the steady state}}} = \underbrace{\mu_i^{ss}}_{\text{standard free energy}} + kT \ln X_i^{O}(cell,t) \qquad (2)$$

The equations above relate the experimental expression level of protein *i* to its value in the steady state:

$$\ln X_i(cell,t) = \ln X_i^O(cell,t) + (\mu_i - \mu_i^O)/kT \quad (3)$$

Given the experimental data $X_i(cell,t)$, which is the measured copy number of protein *i* in a given cell at time, we seek to obtain $X_i^O(cell,t)$, which is the functional protein expression expected at the steady state, as well as $(\mu_i - \mu_i^O)/kT$, which is the deviation from the steady state. Equation (3) represents the changes in the chemical potential for each measured protein. From it we can get the change in the *free energy* of the system as a whole when the constraint(s) are relaxed, ($\Delta G = \sum_i X_i\ (cell,t)(\mu_i - \mu_i^O) \neq 0$). To calculate the steady state expression levels, $X_i^O(cell,t)$, and deviations thereof, $(\mu_i - \mu_i^O)/kT$, we use a procedure described in [9]. Briefly surprisal analysis equates a representation of the data as a sum of terms (right hand side of Equation (1) in the main text) to the logarithm of the measured expression level of protein *i* in the cell c at the time *t*, $X_i(cell,t)$. This is repeated for every cell at every time point *t*. SVD is used as a mathematical tool to determine the two sets of parameters that are needed in surprisal analysis: the cell dependent weights of the constraints $\lambda_\alpha(cell,t)$ (Lagrange multipliers), and the participation of each individual protein *i* in the specific constraint, $G_{i\alpha}$. Determination of the errors for the calculated parameters was performed as described previously.[13] In the numerical procedure we represent the steady state level as $X_i^O(cell,t) = \exp\left(-G_{i0}\lambda_0(cell,t)\right)$.[6, 9] $\lambda_0(cell,t)$ represents the weight of the steady state term in every measured cell. The weight of the proteins when all cellular processes are balanced are described by G_{i0}. In the present study we analysed the expression levels of phosphorylated (activated) enzymes and other functional proteins, the levels of which are expected to be sensitive to the carcinogen perturbations, and thus change with time. The deviations from the steady state are the terms labelled as $\alpha = 1,2$ in the order of their decreasing weight as given by $\lambda_\alpha(cell,t)$. The contribution of each protein to a given unbalanced process λ_α, is given by $G_{i\alpha}$. Thus a change in the chemical potential of protein *i*, $(\mu_i - \mu_i^O)/kT$, due to the constraints $\alpha = 1,2\ldots$ is represented by $\sum_{\alpha=1} G_{i\alpha}\lambda_\alpha(cell,t)$. [7-9] When the steady state term and all the constraints are kept in the exponent in Eq. (1) in the main text, it provides a numerically exact representation of the data and is

not an approximation. Typically however the sum over α in the exponent is truncated to include only the few significant terms. For more details see [7-9].

ST12. Identification of the phase co-existence using Surprisal analysis.

Surprisal analysis can, in principle, use the measured intensities of 11 proteins to identify up to 10 constraints that deviate the system from its stable state. The analysis shows that only the first two constraints $\alpha = 1,2$ are significant, and the second significant process, $\alpha = 2$, contributes only to day 8 (Figure 3a). The transition as seen most clearly in days 12 and 21, is dominated by the first process, $\alpha = 1$ (Figure 3b). To unmask any possible cell individuality, we show the weight $\lambda_1(cell,t)$ of the process $\alpha = 1$ in every *cell* at each measured time point *t*(Figure 3a). To further emphasize the change in value of $\lambda_1(cell,t)$ due to the transition, scatter plots of $\lambda_0(cell,t)$ vs. $\lambda_1(cell,t)$ for days 12 and 28 are shown (Figure 3c). It shows that the two cell populations differ markedly in the values of $\lambda_1(cell,t)$. A key point of surprisal analysis is that the weights of the constraints, the $\lambda_\alpha(cell,t)$'s, are common to all the 11 proteins. However, the weight of the contribution of each protein to any constraint $\alpha = 1,2,..$ is different (Figure S7). $G_{i\alpha}$ is the weight (= importance) of protein *i* in the process α, identifying which proteins are most influencing/most influenced by this constraint. Surprisal analysis determines $G_{i\alpha}$ to have the same value at all times. The change in the protein levels is dictated by the weights of the processes $\alpha = 1,2...,$.

ST13. The response function $\boldsymbol{\Sigma}$:

Using entropy not as a statistical measure of dispersion but as a physical quantity, we apply an approach using the covariance matrix as a response function. For weak perturbations this provides an extension of the well-known qualitative principle of Le Chatelier[2] to characterize protein fluctuations. For strong perturbations this provides the signature of a phase transition in the course of carcinogen treatment. The theory is summarized by the matrix equation $\Delta\bar{\mathbf{N}} = \beta\,\Sigma\,\Delta\boldsymbol{\mu}$. Here, $\Delta\bar{\mathbf{N}}$ is a column vector with 11 components representing change in the average protein levels of the 11 assayed proteins. β is 1/kT, where *k* is Boltzmann's constant, and *T* is temperature. $\boldsymbol{\Sigma}$ is a 11x11 matrix of the experimentally measured covariance of a specific protein P_i with another specific protein P_j. μ is a column vector whose 11 components describe the change in the chemical potentials of 11 proteins, due to a change in external conditions (carcinogen perturbation). To characterize the processes of phase transition we perform diagonalization of the $\boldsymbol{\Sigma}$ at every time

point *t* and compare primary eigenvalues between different time points to identify a time interval where phase transition occurs.

2. Supplementary Figures:

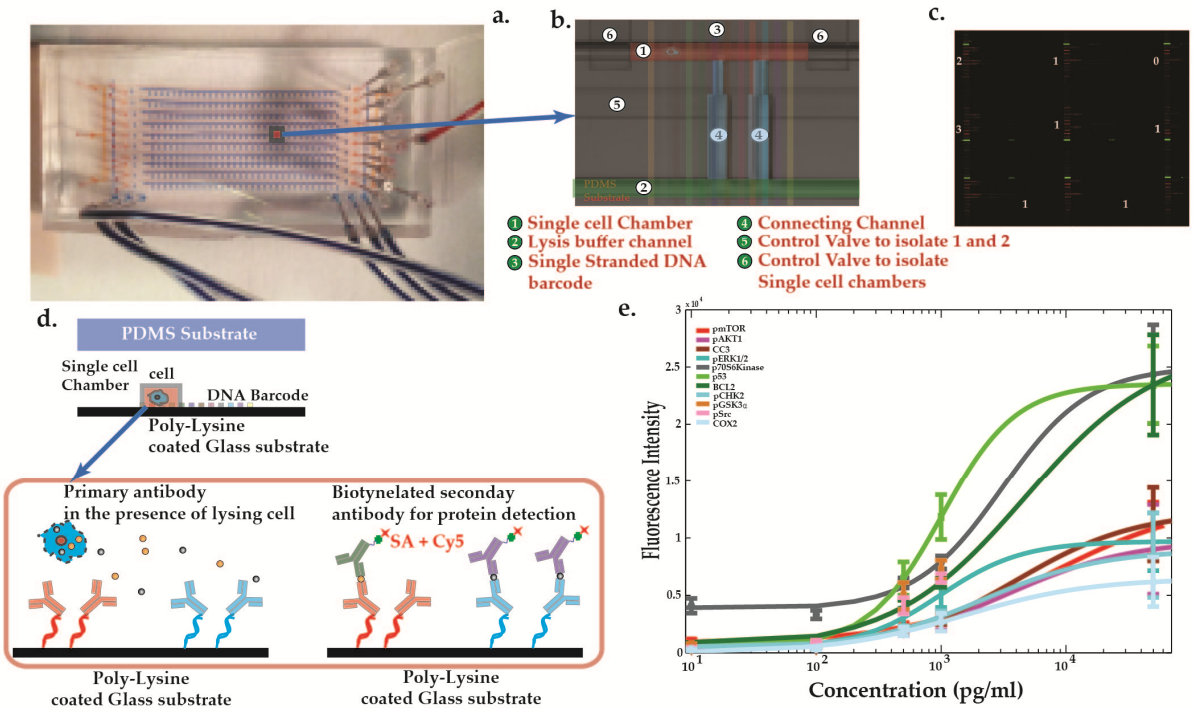

Figure S1: ***Details of the SCBC microfluidic platform.*** *a. shows the overall design layout of the of the SCBC chip, red channels corresponds to the flow layer and the blue layer to the control channels, which is used to manipulate the flow layer. SCBCs has about 300 microchambers, which contains approximately 30% single cell and 25% zero cell chambers (for background normalization). b. Expanded view of a single-cell chamber in the SCBC device. Details of the cell chamber structure is represented by the labelled numerals. Cell/s are trapped in a single channel of the flow channel (1, in figure). c. A sample fluorescence readout of the SCBC slide after the completion of the immune-sandwich assay using Genepix 4400A microarray scanner. Red signals corresponds the actual protein measurement and green signals corresponds to the marker. Numbers inside the image represents the number of cells that were trapped in the respective single cell chambers. d. descriptive diagram of the immuno-sandwich assay used in the SCBC chip for the proteome quantification. e. Calibration plot obtained by measuring the fluorescence intensity of different proteins (inset label in the plot) at different dilution level. Recombinant proteins were used for the measurements. The plot has a sigmoidal shape, similar to the saturation kinetics observed in traditional ELISA assay.*

.

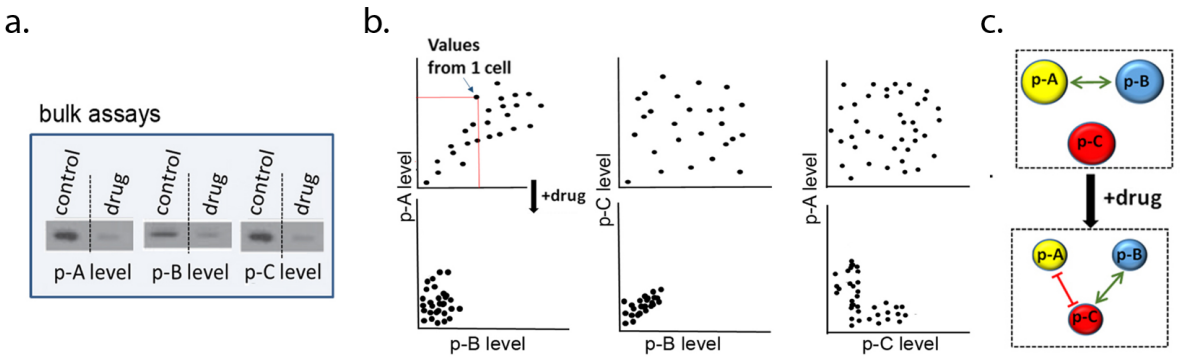

Figure S2: ***Significance of single cell measurements****. Measurements of multiple analytes from the same single cells yield quantitative analyte-analyte correlations (and anti-correlations). This is a uniquely single cell measurement. a. Consider the levels of the three hypothetical phosphoproteins (p-A, p-B, and p-C) intended to represent a small signaling network within a cell. Stimulation (or drugging) of the cell may collectively repress all protein levels as shown in the bulk immunoprecipitation assays. b. Single cell analysis presented in the two-dimensional scatter plots reveals a deeper picture. Note that in the plots for the undrugged cells, all phosphoprotein levels are high, but only p-A and p-B are strongly correlated. Upon drugging, all phosphoproteins are repressed, but p-A and p-B are non-correlated, p-C and p-B are strongly correlated, and p-A and p-B are anti-correlated. c. This inferred correlation network is shown in the network graphic, in which the protein levels are indicated by the sizes of the spheres, and the correlations are indicated by the edges. Correlation, of course, does not mean causation, but a correlation network generated at the single cell level can provide a rich set of testable hypotheses.*

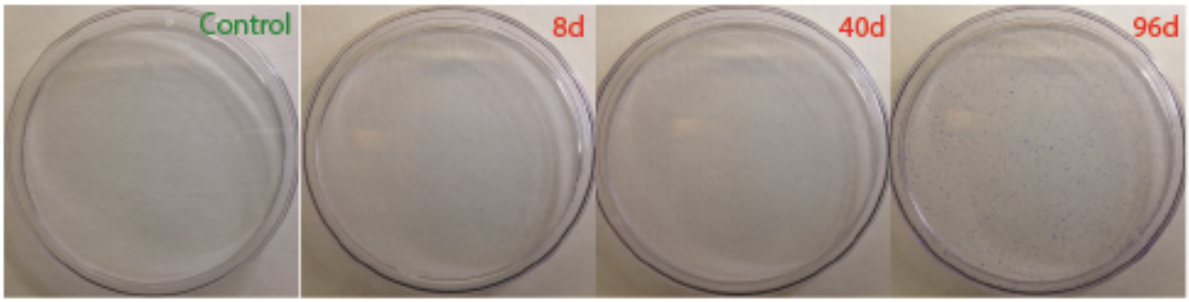

Figure S3: ***Phenotypic characterization of increase in proliferative capacity of carcinogen treated MCF-10F as a function of time****. Results of colonogenic assay of cells treated with BaP, to determine dependence on mitotic growth factor. The cell were grown in medium having reduced growth factor. At day 96, there is a stark increase in proliferative capacity of cells treated with carcinogen in low growth factor medium.*

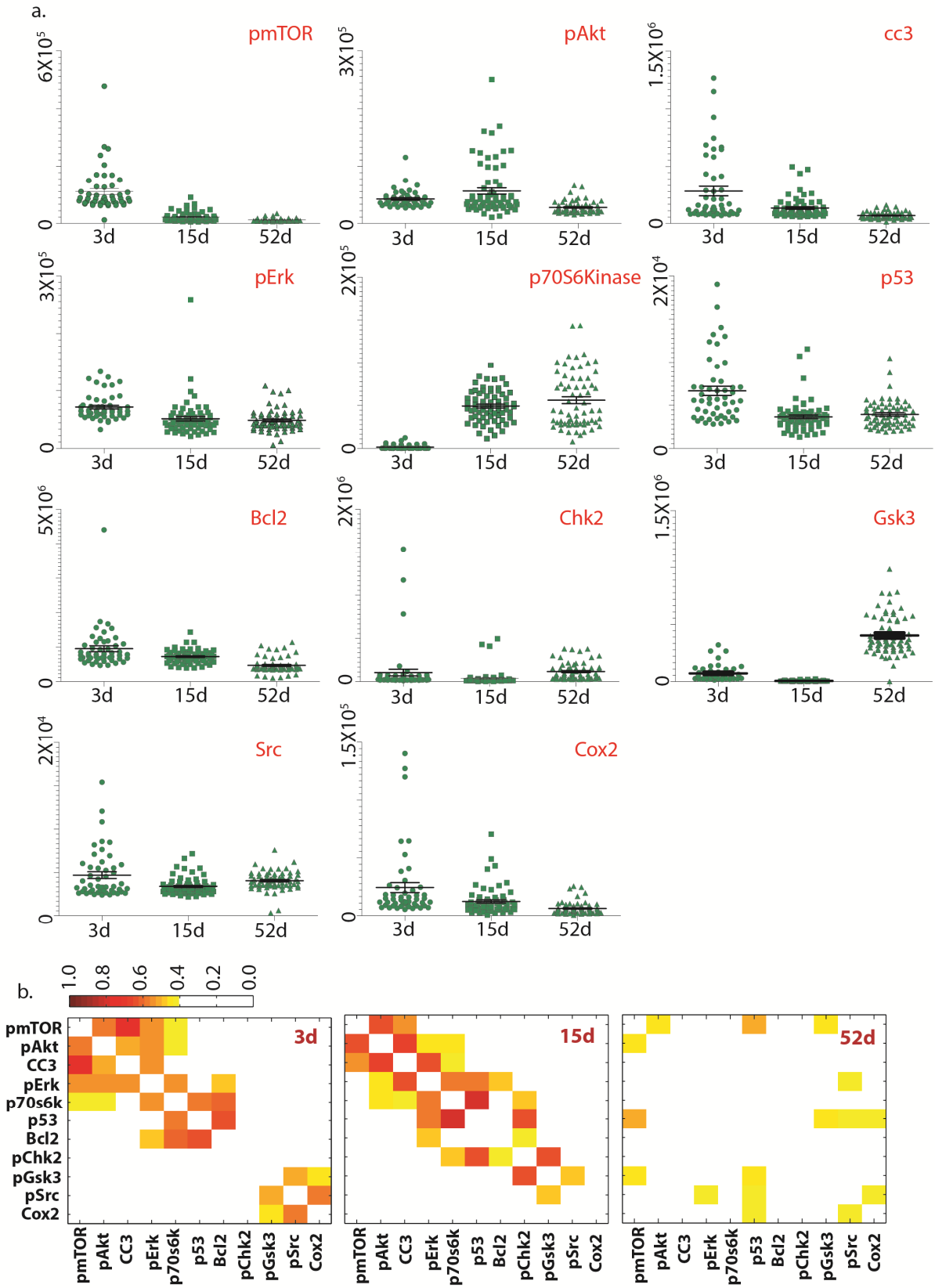

Figure S4: ***Results of SCBC analysis of control samples****. a. Protein-protein covariance matrices, extracted from SCBC data of MCF-10F control cells passaged regularly, at three different time points.*

b. One-dimensional scatter plots of the single cell levels of all the proteins in the panel, at the corresponding time points of the control samples.

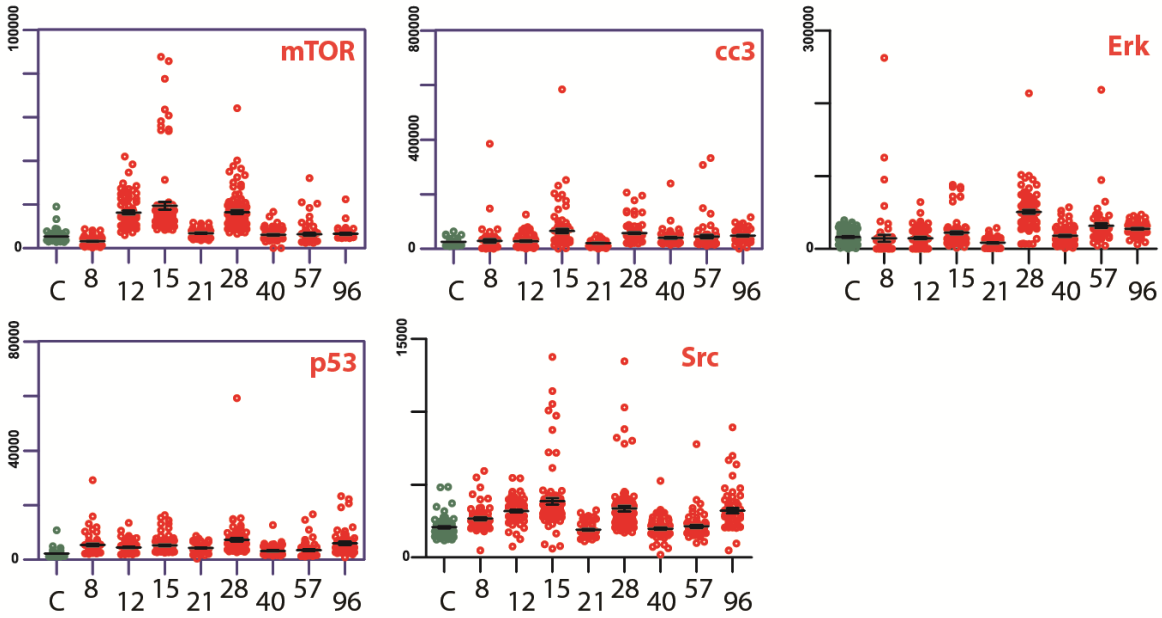

Figure S5: ***Scatter plot of protein panels analysed using the SCBC platform.*** *Variations (distribution) represented by 1D scatter plot for the remaining proteins used in this work, measured at different time points during and post treatment of MCF-10F cells with BaP. There is a significant increase of variations in copy numbers for most proteins during the transition window of 8-28 days*

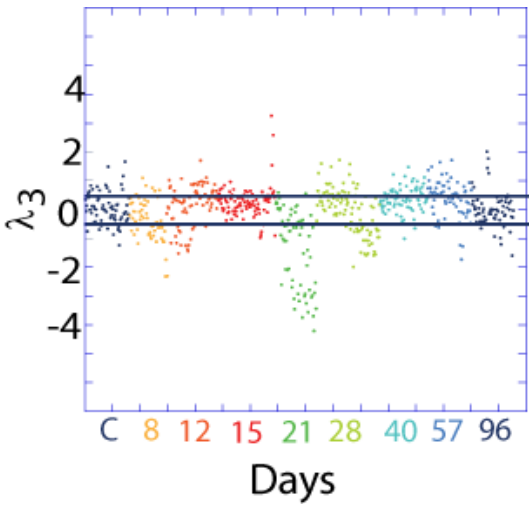

Figure S6: ***Constraint*** $\alpha = 3$ ***in B[a]P treated cells.*** *Constraint* $\alpha = 3$ *is identified using Surprisal analysis and its amplitude,* $\lambda_3(c,t)$*, is shown for all single cells as a function of time. The black lines are error limits due to noise in the data.*

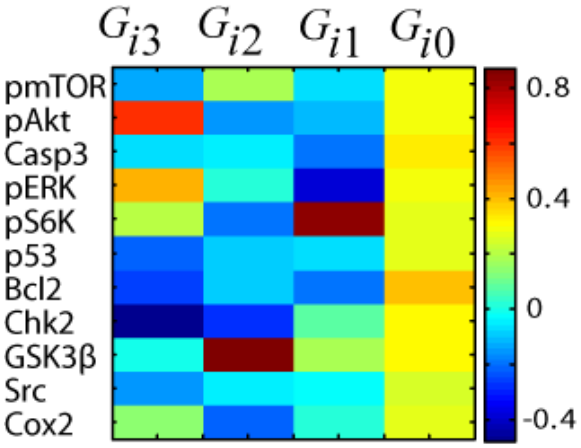

Figure S7: ***Extents of participation of the assayed proteins in the steady state and unbalanced processes****. Surprisal analysis yields the extent of participation of each assayed protein,* $G_{i\alpha}$*, in the biological processes described by constraints* $\alpha = 0$ *(the steady state) and* $\alpha = 1,2,3$ *in the B[a]P treated cell population.*

3. Supplementary Table:

Table S1: List of DNA oligoneucleotides and Antibody pairs used for the SCBC Immunoassays used in this work. All oligonucleotides were synthesized by Integrated DNA Technology (IDT) and purified via high performance liquid chromatography (HPLC).

DNA	DNA Sequence	
C	AAAAAAAAAAAAAGCACTCGTCTACTATCGCTA	
C'	/5AmMC6/AAAAAAAAAAAAATAGCGATAGTAGACGAGTGC	
D	AAAAAAAAAAAAAATGGTCGAGATGTCAGAGTA	
D'	/5AmMC6/AAAAAAAAAAAAATACTCTGACATCTCGACCAT	
E	AAAAAAAAAAAAAATGTGAAGTGGCAGTATCTA	
E'	/5AmMC6/AAAAAAAAAAAAATAGATACTGCCACTTCACAT	
F	AAAAAAAAAAAAAATCAGGTAAGGTTCACGGTA	
F'	/5AmMC6/AAAAAAAAAAAAATACCGTGAACCTTACCTGAT	
G	AAAAAAAAAAAAAGAGTAGCCTTCCCGAGCATT	
G'	/5AmMC6/AAAAAAAAAAAAAAATGCTCGGGAAGGCTACTC	
H	AAAAAAAAAAAAAATTGACCAAACTGCGGTGCG	
H'	/5AmMC6/AAAAAAAAAAAAAACGCACCGCAGTTTGGTCAAT	
I	AAAAAAAAAAAAATGCCCTATTGTTGCGTCGGA	
I'	/5AmMC6/AAAAAAAAAAAAATCCGACGCAACAATAGGGCA	
J	AAAAAAAAAAAAATCTTCTAGTTGTCGAGCAGG	
J'	/5AmMC6/AAAAAAAAAAAAAACCTGCTCGACAACTAGAAGA	
K	AAAAAAAAAAAAATAATCTAATTCTGGTCGCGG	
K'	/5AmMC6/AAAAAAAAAAAAAACCGCGACCAGAATTAGATTA	
L	AAAAAAAAAAAAAGTGATTAAGTCTGCTTCGGC	
L'	/5AmMC6/AAAAAAAAAAAAAGCCGAAGCAGACTTAATCAC	
M	AAAAAAAAAAAAAGTCGAGGATTCTGAACCTGT	
M'-Cy3	/5Cy3/AAAAAAAAAAAAAACAGGTTCAGAATCCTCGAC	
N	AAAAAAAAAAAAAGTCCTCGCTTCGTCTATGAG	
N'	/5AmMC6/AAAAAAAAAAAAAACTCATAGACGAAGCGAGGAC	

DNA Label	Antibody Kit	Catalogue Number
C'	Human Phospho-TOR (S2448) DuoSet ELISA Kit (R&D systems)	DYC1665
D'	Human/Mouse Phospho-Akt1 (S473) DuoSet ELISA Kit (R&D systems)	DYC2289C
E'	Human/Mouse Cleaved Caspase-3 (Asp175) DuoSet ELISA Kit (R&D systems)	DYC835
F'	Human/Mouse/Rat Phospho-ERK1 (T202/Y204)/ERK2 (T185/Y187) DuoSet ELISA Kit (R&D systems)	DYC1018B
G'	Human Phospho-p70 S6 Kinase (T389) DuoSet ELISA Kit (R&D systems)	DYC896
H'	Human Total p53 DuoSet ELISA Kit (R&D systems)	DYC1043
I'	Human Total Bcl-2 DuoSet ELISA Kit (R&D systems)	DYC827B
J'	Human Phospho-Chk2 (T68) DuoSet ELISA Kit (R&D systems)	DYC1626
K'	Human/Mouse/Rat Phospho-GSK-3 alpha/beta (S21/S9) DuoSet ELISA Kit (R&D systems)	DYC2630
L'	Human Phospho-Src (Y419) DuoSet ELISA Kit (R&D systems)	DYC2685
N'	Human/Mouse Total COX-2 DuoSet ELISA Kit (R&D systems)	DYC4198

1. Shi, Q.; Qin, L.; Wei, W.; Geng, F.; Fan, R.; Shin, Y. S.; Guo, D.; Hood, L.; Mischel, P. S.; Heath, J. R., *Proc Natl Acad Sci U S A* **2012,** *109* (2), 419-24. DOI 10.1073/pnas.1110865109.
2. Shin, Y. S.; Remacle, F.; Fan, R.; Hwang, K.; Wei, W.; Ahmad, H.; Levine, R. D.; Heath, J. R., *Biophysical journal* **2011,** *100* (10), 2378-86. DOI 10.1016/j.bpj.2011.04.025.
3. Calaf, G.; Russo, J., *Carcinogenesis* **1993,** *14* (3), 483-92.
4. Franken, N. A.; Rodermond, H. M.; Stap, J.; Haveman, J.; van Bree, C., *Nat Protoc* **2006,** *1* (5), 2315-9. DOI 10.1038/nprot.2006.339.
5. Siriwardhana, N.; Choudhary, S.; Wang, H. C., *Breast Cancer Res Treat* **2008,** *109* (3), 427-41. DOI 10.1007/s10549-007-9666-9.
6. Kravchenko-Balasha, N.; Levitzki, A.; Goldstein, A.; Rotter, V.; Gross, A.; Remacle, F.; Levine, R. D., *Proc Natl Acad Sci U S A* **2012,** *109* (12), 4702-7. DOI 10.1073/pnas.1200790109.
7. Kravchenko-Balasha, N.; Remacle, F.; Gross, A.; Rotter, V.; Levitzki, A.; Levine, R. D., *BMC systems biology* **2011,** *5*, 42. DOI 10.1186/1752-0509-5-42.
8. Kravchenko-Balasha, N.; Wang, J.; Remacle, F.; Levine, R. D.; Heath, J. R., *Proceedings of the National Academy of Sciences of the United States of America* **2014,** *111* (17), 6521-6. DOI 10.1073/pnas.1404462111.

9. Remacle, F.; Kravchenko-Balasha, N.; Levitzki, A.; Levine, R. D., *Proc Natl Acad Sci U S A* **2010,** *107* (22), 10324-9. DOI 10.1073/pnas.1005283107.
10. Zadran, S.; Remacle, F.; Levine, R. D., *Proceedings of the National Academy of Sciences of the United States of America* **2013,** *110* (47), 19160-5. DOI 10.1073/pnas.1316991110.
11. David L. Nelson; Albert L. Lehninger; Cox, M. M., *Lehninger Principles of Biochemistry*. 5 ed.; Macmillan: **2008**.
12. Mayer, J. E.; Mayer, M. G., *Statistical mechanics*. 2nd ed.; John Wiley&Sons: New York, **1977**.
13. Gross, A.; Levine, R. D., *PLoS One* **2013,** *8* (4), e61554. DOI 10.1371/journal.pone.0061554.